‘Placebos’ and the Logic of Placebo Comparison

Andrew Turner

Abstract

Robin Nunn has argued that we should stop using the terms ‘placebo’ and ‘placebo effect’. I argue in support of Nunn’s position by considering the logic of why we perform placebo comparisons. Like all comparisons, placebo comparison is just a case of comparing one thing with another, but it is a mistake, I argue, to think of placebo comparison as a case where something is compared to ‘a placebo’. Rather, placebo comparison should be understood as a situation which sets-up the treatment and control groups in a particular way; not as a case involving objects or procedures called ‘placebos’ employed in order to control for ‘placebo effects’.

Keywords

Placebo, Placebo controlled trial, Clinical trials, Evidence-based medicine.

‘Placebos’ and the Logic of Placebo Comparison

Robin Nunn has argued that we should stop using the terms ‘placebo’ and ‘placebo effect’ (Nunn 2009a, b). He claims the terms are fraught with conceptual confusion, and that there is good empirical evidence that lumping a disparate range of elements together under these terms is, in essence, like mixing paint colours to get brown (See the two meta-analyses: Hrobjartsson and Gøtzsche (2001, 2004)). The point being that, if we do wish to say something informative about medical treatments, ‘placebo’ and ‘placebo effect’ are not terms which are analytically useful[1]. Instead, we should always be much more specific about the particular details of particular therapeutic situations; and when we are, we stop needing to use the term ‘placebo’ or ‘placebo effect’.

I agree. In what follows I argue in support of Nunn’s position. Importantly I think that much of the work needed to support Nunn’s position can be achieved through considering the logic of why we perform placebo comparisons. Like all comparisons, placebo comparison is just a case of comparing one thing with another, but it is a mistake, I argue, to think of placebo comparison as a case where something is compared to ‘a placebo’. What we compare are treatment groups, or better, the average effect sizes in our groups. Placebo comparison should be understood as a situation which sets-up those groups in a particular way; not as a case involving objects or procedures called ‘placebos’ employed in order to control for ‘placebo effects’.

In essence my argument is an elaboration of a simple idea, which is neatly summed up by Austin Bradford Hill:

> ‘To some patients a specific drug is given, to others it is not. The progress and prognosis of these patients are then compared. But in making this comparison in relation to the treatment the fundamental assumption is made – and must be made – that the two groups are equivalent in all respects, except for the difference in treatment’ (Hill 1951: 278)

[1] The concern in this paper is with placebo comparison, and so the argument is about the use of the term ‘placebo’ in a research context. Clearly however, the term has uses in clinical contexts that may be valid for independent reasons. I would expect the argument made here to apply to the clinical context; since the general idea is simply that the term obscures what can be better explained in more precise terms. However I will not argue explicitly for this here.

The logic of placebo comparison

Consider first some preliminary points about the purpose of placebo comparison. I claim that the key *epistemic* aim of placebo comparison, which is what is important to this discussion, is to learn about the efficacy of particular aspects of a treatment. That is not to say that there might not be other aims in mind when placebo comparisons, or placebo controlled trials (PCTs), are performed – such as having to meet regulatory requirements on the road to getting a new treatment approved, or performing a trial that is more likely to show a new treatment in a positive light (as opposed, for example, to comparing with the current best treatment). These other, more instrumental, aims will not be the focus of my argument however.

We should note that while I claim the aim of placebo comparison is to learn about efficacy, placebo comparison is not, by any means, the *only* way we can learn about efficacy. However it is, prima facie, a good way. Indeed a principle such as 'a treatment is efficacious if and only if it outperforms placebo' looks very tempting[2]: it underlies, for example, the often rehearsed argument that PCTs possess unparalleled 'assay sensitivity' (Temple and Ellenberg 2000), though whether placebo comparisons really do possess significant epistemic virtue, over and above other comparisons, has recently been questioned (Howick 2009). But in the ideal case at least, the logic of placebo comparison is well-equipped to give us insight into the efficacy of the aspect of the treatment under investigation.

One might also claim that placebo comparisons give us insights into placebo effects, at least in so far as we take an interest in the outcomes observed in the placebo groups of PCTs. Ernst and Resch (1995) distinguish between, what they call, true and perceived placebo effects. They note that observations of outcomes in placebo groups confound the 'true placebo effects' with other effects that result from natural variations in the condition. So the outcomes of the placebo group in a PCT do not represent placebo effects as such, but only 'perceived placebo effects'. In what follows however, the concern

[2]Although note that such a principle does not imply that we cannot know that a treatment is efficacious unless a placebo comparison has been performed. It simply equates efficacious treatments with placebo outperforming treatments. If we know a treatment is efficacious, that justifies the belief that it would outperform placebo, were such a comparison to be performed.

will be with the comparative nature of PCTs, and not with what may or may not be inferred from the experimental groups individually.

Consider now the logic of placebo comparison. The paradigm case of placebo comparison is the PCT of a drug. Such a comparison is done in order to measure the capacity of the drug contained in the treatment to produce therapeutic effects. To avoid confusion and to make clear what is meant by talking in terms of 'aspects of a treatment' we can stipulate a distinction between drug and treatment. Take 'drug' to denote the (allegedly) therapeutic chemical or chemicals, and take 'treatment' to denote a delivery system, perhaps but not necessarily containing a drug. Hence for clarity I mean to set-up the terms such that drugs are not pills, but treatments can be pills (though of course things besides pills can be treatments), and a pill may or may not contain a drug while still remaining a treatment, etc. We can also widen this definition of treatment to include not just the object which is delivering the drug, but also the way in which it is delivered. So for instance we could talk about the kindness of the healthcare professional, or the patient's feeling of hope, as being some of the *contextual aspects* of a treatment, just as a drug is a *pharmacological aspect* of a treatment. Consequently we can say that *the efficacy of the drug* is the *aspect of the treatment* that we wish to investigate in a *PCT of the drug*.

The logic behind placebo comparison is straightforward, especially when put in terms of trials of drugs. In the ideal case we compare two groups which are identical in all therapeutically relevant respects, but for the fact that one group receives the drug whereas the other group does not. This is precisely the point expressed in the quotation from Austin Bradford Hill, above.

Note that we must compare the presence and absence of a *drug*: the comparison between the presence and absence of a *treatment* is a very different comparison. The point of comparing two groups that differ only in regards to the presence of a drug is that it allows us to infer that any differential effects between the groups can be attributed to the drug's action. This therefore allows us to reasonably claim that the drug caused those differential effects. Indeed in the ideal case this method is, as Nancy Cartwright calls it, a 'clincher', meaning that the causal conclusion is deductively implied (Cartwright 2007).

We can generalise the logic beyond trials of drugs. The efficacy of a drug is only one aspect of a treatment, and there are many different aspects of a treatment that we might wish to investigate the efficacy of, beside its drug content. The logic of placebo comparison is indifferent to whether the particular

aspect to be singled out happens to be a treatment's drug content. For example, consider the following case: We investigate whether a treatment consisting of a pill containing x mg of drug performs better than a treatment consisting of two pills, one of which contains x mg of drug and the other of which is a sugar pill[3]. In that case, it would be the efficacy of 'receiving an extra sugar pill' that would be the aspect of the treatment we were investigating; because that is the aspect of the treatment that has been singled out.

The logic of placebo comparison simply involves singling out particular aspects of treatments, to which we may or may not be able to attribute efficacy. There is no logical requirement to only attribute efficacy to the action of drugs. I claim that this should be uncontroversial: it really is nothing more than an elaboration of the Hill quote above.

Where do 'placebos' enter into the logic of placebo comparison?

The following is (part of) an influential, but much criticised (see for example: Gøtzsche 1994, 1995; Grünbaum 1991, 1981; Moerman 2002; Miller and Kaptchuk 2008), definition of a placebo put forward by Arthur and Elaine Shapiro (see different versions of it in: Shapiro (1964, 1968); Shapiro and Shapiro (1997b)):

> '[A placebo] is any therapy prescribed knowingly or unknowingly by a healer, or used by laymen, for its therapeutic effect on a symptom or disease, but which actually is ineffective or not specifically effective for the symptom or disorder being treated.' (Shapiro and Shapiro 1997a: 12)

Consider that this definition entails that 'patting ones head' might be a placebo for headaches, if a doctor (knowingly or unknowingly) recommended this to you as a supposedly effective treatment. The fact that the doctor recommends it as a treatment for headaches fulfils the first part of the definition. The second part is fulfilled because, as we know, 'patting one's head' is (at least under usual circumstances) actually ineffective for treating headaches: more likely it will make it worse.

[3] Or more precisely, a pill with no therapeutically relevant contents (cf. Golomb 1995; Golomb et al. 2010)

‘Patting one’s head’ however would be entirely useless in a PCT of aspirin, despite the fact that, according to the Shapiros’ definition, it is a placebo headache treatment. The reason it would be useless is clear from the logic of placebo comparison explained above. To reiterate, comparing ‘patting one’s head’ with aspirin is not a comparison which singles out only the effect of the particular aspect of the treatment that is under investigation: namely, the action of the drug aspirin. So even if we take the Shapiros’ definition seriously (and many argue we shouldn’t), the fact that something might, according to that definition, be a ‘placebo’ treatment for X, does not guarantee that it would be useful in a PCT of some other treatment for X.

The reason it is instructive to look at the Shapiros’ definition of a placebo – even though it is highly criticised – is that it embodies an intuitive idea about ‘placebos’. Namely, the idea that ‘placebos’ are particular things, or in other words, that it makes sense to claim that such-and-such is ‘a placebo’. Such an idea is by no means unique to the Shapiros’. The point of the head-patting example aims directly at that intuitive idea, and in that sense generalises beyond the Shapiro’s definition. The underlying assumption of any definition of ‘a placebo’ is that ‘placebos’ are conceptually prior to placebo comparisons: as if it were possible to take a jar of ‘placebos’ off the shelf, ready to use in some forthcoming PCT. I claim that this is false. For any candidate definition of ‘a placebo’ we can find an object that would fill the definition, but imagine a situation in which we compare it with another treatment and fail to produce a comparison, which follows the logic set out above, and which I claim is the logic of *placebo comparison*.

Now it may be argued that the head-patting example only shows that there are such things as bad placebo comparisons; so that the example is, contrary to my suggestion, an example of a placebo comparison (because it involves ‘a placebo’), but a bad one (because it doesn’t follow the logic). Instead I claim that we should not understand placebo comparison, good or bad, as involving the comparison of one thing – a placebo – with another – the ‘active’ treatment. I will argue below that whether one is performing a placebo comparison depends only on whether one follows the logic set out above, and in no way depends on whether the comparison involves particular objects or procedures that some may call ‘placebos’.

What counts as the 'placebo group' depends entirely on the intended comparison

Branthwaite and Cooper (1981) investigated the therapeutic effect of branded packaging. They made a four way comparison of branded and unbranded, aspirin and sugar pills[4]. They found that branded packaging consistently provided more relief of headaches: "Branding appeared to supplement both the inert placebo and the active ingredients to produce more relief than either placebo or active ingredients alone" (1981: 1578). Their result however it not the focus here, rather it is the fact that in their paper Branthwaite and Cooper call the branded and unbranded sugar pills 'placebos' and the groups which received these sugar pills the 'placebo groups'.

This is an intuitive way to label the groups if we think that 'placebos' are things, since such labelling follows straightforwardly from the 'sugar pill = a placebo' idea. Sugar pills are placebos, groups given sugar pills are, therefore, placebo groups.

I claim that Branthwaite and Cooper have labelled their groups incorrectly. More precisely I claim that which of their groups we choose to call the placebo group is, without further specification, undetermined. The reason is that, as I suggested above, the placebo group identifies a group playing a particular logical role in a comparison; namely, keeping all but one of the therapeutically relevant aspects of the treatment identical. From Branthwaite and Cooper's four groups we can make a number of different comparisons, and it is only with specific reference to some particular comparison that it makes sense to invoke the term placebo group.

So: If we're interested the differential effects due to branding between the two groups receiving aspirin containing pills, then the placebo group in that case would be the group receiving the non-branded aspirin pills. And if we're interested the differential effects due to branding between the two groups receiving sugar pills, then the placebo group would be the group receiving the non-branded sugar pill. If we were interested in the differential effects due to aspirin in the two branded groups, the placebo group in that case would be group

[4]Actually I'm just assuming they were *sugar* pills. We are told in the Branthwaite & Cooper's methods section ([1981] p.1576) that the pills not containing aspirin were the same size, shape, weight and colour, and that they were not designed to taste the same as aspirin tablets. The content of these pills is not disclosed.

receiving the branded sugar pill. And lastly, if we were interested in the differential effects due to aspirin in the two unbranded groups, the placebo group would be group receiving the unbranded sugar pill.

Equally it would make no sense, for example, to call the group receiving the unbranded sugar pill a placebo group when compared to the branded aspirin group, because in that case more than one aspect of the treatment is being singled out, and to reiterate, the logic behind placebo comparison is to single out only one particular aspect of a treatment. My criticism of Branthwaite and Cooper's labelling of their groups is simply that we can pick a number of different pairs (four pairs, in fact) of their four groups which are identical in all but one respect (as enumerated above). So any particular group may or may not be labelled a 'placebo group' depending on which pair of groups we have in mind, or in other words, depending on what comparison is being made. The general point that this enumeration labours is that we shouldn't call one group a placebo group, independently of a particular comparison.

This is certainly not a re-labelling a medical researcher would likely endorse, and there is a clear objection to consider here. It is an objection to my claim that as long as the two groups being compared are identical in all but one respect, then we have a placebo comparison - Isn't it just wrong to claim this? Won't any sensible medical researcher object that a comparison, say, of branded versus unbranded aspirin is no more a 'placebo comparison' than a comparison between 5mg and 10mg of a drug: these are more properly called 'active' comparisons.

Such an objection would seem to rest on the known 'activity' of aspirin, namely the fact that aspirin pills contain a chemical (2-acetoxybenzoic acid) with a well understood analgesic effect, whereas placebos are not thought of as containing pharmacologically relevant chemicals. So, if two groups were to receive aspirin-containing pills, and those groups differ only in respect of whether or not the pills were branded, then neither group has received a placebo. Therefore, it is not a placebo comparison; despite what I might choose call the underlying logic of that comparison. The objection relies on the idea that we can distinguish placebo from non-placebo aspects of a treatment by their mechanism. The active aspects, like aspirin-content, work through a known chemical and biological mechanism, and, so the argument goes, the placebo aspects work through placebo mechanisms that are relevantly different enough to justify making a distinction between active and placebo comparisons. More

needs to be said about ‘placebo mechanisms’ for the objection to be convincing, and indeed research into – so called – ‘placebo effects’ may seem to provide some apparent support.

Price et al (2008) note in their review of the placebo-research literature that recent work has conceptualised ‘placebo effects’ in terms of ‘the psychosocial context surrounding the patient and the effect that this context has on the patient’s experience, brain, and body’ (2008: 567). Others have captured this idea by recognising that what is important is the ability of objects or procedures to generate therapeutic responses in virtue of the meaning or the symbolism that they have for the patient. Indeed such meaning-theories are prominent in the placebo literature (Moerman et al. 1979; Moerman 2002; Moerman and Jonas 2002; Thompson et al. 2009). Consequently the most interesting, and perhaps most coherent, approach to understanding placebo effects suggests that they should be conceived of as the result of a range of context-specific psychological and social factors, operating through specific physiological mechanisms (Hahn and Kleinman 1983; Papakostas and Daras 2001; Kaptchuk 2002; Kirmayer 2004; Moerman 2002; Moerman and Jonas 2002; Price 1984; Stein 1983).

The point therefore is that placebo mechanisms have in common the fact that they involve a response to the meaning of some aspect(s) of a treatment; which, as the objector to my claims will note, is a fact that provides sufficient basis to distinguish between responses generated in those ways, and responses generated by, for example, pharmacological content. Placebo comparisons, as anyone pressing the objection would reiterate, are those comparisons where the observed effects in one or both groups are generated in response to the meaning of the treatment. Sugar pills are called placebos, because the only conceivable way they could have a therapeutic effect is through these meaning-based placebo mechanisms. A comparison between 10mg and 15mg of a drug is just not that kind of comparison, and at most, a comparison of branded and unbranded aspirin-containing pills could be thought of as involving a pharmacologically enhanced placebo.

I claim this view is not tenable. If we characterise placebo comparisons by the presence of objects or procedures which are generating their therapeutic effects in virtue of the meaning attached to them, then (as illustrated in the head patting example above) the simple fact that one is performing a placebo comparison, in that sense, need not entail that one is following the logic set out

above. Instead our objector now needs to ask of any given placebo comparison (given it is supposed to be one which measures the efficacy of some aspect of a treatment) whether it does follow the logic set out above. Hence on the objector's view, placebo comparison (that is, comparison with 'a placebo') and efficacy testing need have no connection to each other. Rather, the objector's view is that 'placebos' are just another category of objects and procedures, which may be called upon as a control in a clinical trial that may or may not be investigating the efficacy of some aspect of a treatment. The key question to ask is what work the distinction between placebo and non-placebo comparisons is supposed to do here – given that it is not to do with efficacy testing.

I suggest that this division of comparisons into placebo and non-placebo is an arbitrary division to make. It is certainly not made on the basis that comparison with placebos allows us to attribute efficacy to aspects of a treatment, whereas comparison with 'active' treatments does not. Because as set out above, on the objector's view, whether a comparison is with 'a placebo' or not has nothing to do with whether it is aim is to measure efficacy.

We could, equally well, stipulate to divide comparisons into those involving treatments with an aspect that works through the renin-angiotensin system (e.g. the ACE inhibitors - ramipril etc). That distinction too has nothing to do with efficacy testing, and it too divides comparisons according to the mechanism by which therapeutic responses are generated. The point is that it serves no useful analytical purpose to divide our clinical trials into those featuring controls that work through the renin-angiotensin system and those that do not, based on the presence or absence of, for example, ACE inhibitors in the trial. Just as, I claim, it serves no useful analytical purpose to call a highly heterogeneous set of objects and procedures 'placebos' and to divided our clinical trials into placebo and non-placebo controlled, based on the presence or absence of such objects in the trial. And given the diversity of biopsychosocial factors and mechanisms that 'placebo effects' are suppose to encompass, a division based on the variety of placebo mechanisms is a great deal less clear than a division based on the renin-angiotensin system.

Contrary to this confusing characterisation of 'placebos' I suggest that, if we think carefully about the logic of placebo comparison, then we don't need to talk about 'placebos' at all.

Placebo comparison without 'placebos'

Placebo comparisons are those which compare two groups that are identical in all but one respect. How this identity is ensured, or approximated to, is a question of trial design. The placebo group in a PCT needs to be designed so as to ensure the required identity, and as illustrated above, a group which is told to ‘pat one’s head’ is certainly not a legitimate placebo group for a PCT of the drug aspirin. But, fairly obviously, a group given sugar pills exactly like the aspirin-containing pills has much more potential to be a legitimate placebo group in a PCT of aspirin. The question of what objects or procedures are necessary for any particular PCT depends on the nature of the treatment as a whole, and the aspect of that treatment which is being investigated.

The common equation of ‘placebos’ with sugar pills is the result of the fact that pills are a paradigmatic example of a drug delivery system. It is almost too obvious to state that if a treatment is in the form of a single pill containing a drug, then it makes sense to give patients in the placebo group an exactly similar non-drug-containing pill in order to avoid confounding the therapeutic action of the drug with the therapeutic action of simply giving a pill. The fact that this is so obvious makes it possible to underrate its significance. It tempts us to make the mistake of trying to identify placebos with sugar pills, rather than taking the correct view that, across many circumstances, sugar pills are merely highly apt to ensure identity between the treatment and placebo groups with respect to ‘receiving a pill’. We sometimes give sugar pills to a placebo group to meet the requirement of keeping treatment groups identical in all but one therapeutically relevant respect: we do not do it because those pills are ‘placebos’.

The reiterate: sugar pills are not a special kind of object called ‘placebos’; it just happens that sugar pills are a particularly easy to grasp example of an object that might do the work of controlling for certain therapeutically relevant aspects of a treatment, when we perform a placebo comparison. There is no such thing as ‘a placebo’, but there are certain ‘control roles’ that need be played in placebo comparisons, just as in any meaningful comparison. If placebo comparisons are a special kind of comparison, it is not because they involve comparison with a special kind of object (‘a placebo’), but because they involve a control group (the placebo group) with special features. And we know precisely what those special features are: they are those that ensure the placebo group is identical to the treatment group in all but one respect.

Now it might be argued that ‘placebo’ is simply a shorthand way of labelling an experimental control such as an ‘exactly similar non-drug-containing

pill', or perhaps in suitably different circumstances, the act of 'patting one's head'. This would be an argument for the view that the notion of 'a placebo' does in fact make sense, when restricted to the context of some particular comparison and on the understanding that 'a placebo' in one context may not remain 'a placebo' in another. Hence, relative to a particular comparison we could quite legitimately point to some object or procedure and call it 'a placebo'. Such a view asserts that the term 'placebo' is not meaningless or unhelpful. On the contrary it purports to do the helpful work of summing up important details about the control being used in a given comparative trial; and neither does it involve distinguishing objects and procedures on any mechanistic basis.

This 'placebo-shorthand' view does not succeed however. Our two groups should be identical in all therapeutically relevant respects, except the one under investigation. And we have said above that a placebo group is a group which possesses the specific features which ensure this identity. Now ask, what is the term 'placebo' supposed to go shorthand for? – Presumably, it should go shorthand for a set of measures we have taken to ensure the identity, between groups, of some of the therapeutically relevant aspects of the treatment: but what set? – If we mean the measures taken to ensure the identity of *all* the therapeutically relevant aspects, besides the one being investigated, then we already have a name for that, that is just *the placebo group*. Of some purported placebo group, we need to know whether it genuinely does possess the features that would enable a legitimate placebo comparison. That consists of asking questions about particular aspects of the treatment, such as whether the delivery mechanisms are the same, whether the patients are given the same information, whether the doctors have the same expectations for the two groups etc, and importantly there are no questions, at this level of specificity, that involve talking about 'placebos'.

If however we stipulate that the term 'placebo' should go shorthand for some proper subset of measures, then that fails to be helpful. Since we still need to ask the same, more specific, questions about each element in the shorthand account; in order to assess whether the purported 'placebo' is legitimate. And moreover the knowledge that that proper subset of measures genuinely ensures the required identity between *only some* therapeutically relevant aspects of the treatment, still does not guarantee the legitimacy of the placebo group as a whole: since the legitimacy of the placebo group depends on *all* aspects (but the one under investigation) being identical between groups. So for example, if we

are conducting a PCT of the drug aspirin, delivered in pill form, we could choose to call our exactly similar non-drug-containing pills ‘placebo pills’. But we could still fail to conduct a legitimate placebo comparison with these ‘placebo pills’; perhaps because our two groups were, say, given very different information and reassurance as part of their respective treatments. And just because we had called them placebo pills, that would not remove the need to ask specifically, whether they were similarly coloured, shaped, or possessed no relevantly-active content – which is what we would have to do anyway, even if we hadn’t called them placebo pills. The placebo-shorthand view fails because it has no bearing on those questions we have to ask of any placebo comparison, to ensure it is a good one. We could certainly stipulate to call certain kinds of control measures ‘placebos’ as a shorthand, but only because we can make any number of such redundant shorthand stipulations – and even if we were to do this, it would still be a stipulation that only made sense with respect to the particular comparison being performed.

The key point is that it is the specific details of the placebo group, as a whole, that matter for placebo comparison. The fact that we could stipulate that a certain subset of features of a particular placebo group should be called ‘a placebo’ does not solve any problems. It is redundant to call anything ‘a placebo’, even with respect to some particular comparison.

Meaning-theories of placebo demonstrate which aspects of a treatment might be therapeutically relevant

I argued above that meaning-theories of placebo do not enable us to make a useful distinction between ‘placebo’ and ‘active’ comparisons. Never the less many of the empirical results which inform meaning-theories have taken much of the mystery out of – what some would call – ‘placebo effects’ or ‘placebo responses’. The fact that we can give sophisticated empirical accounts of the psychological and physiological mechanisms by which expectations, beliefs, desires etc have therapeutic effect does not necessitate using the term ‘placebo effect’, ‘placebo response’, ‘meaning response’ or any of the other candidate re-phrasings (Nunn 2009a). What these empirical results do highlight however are the many different aspects of a treatment that can be therapeutically relevant.

This is significant because, while the logic of placebo comparison tells us to keep therapeutically relevant aspects of a treatment identical between

groups, it obviously does not tell us which aspects are therapeutically relevant[5]. The therapeutic effects of many different aspects of a treatment have been investigated. The surprising result – which the meaning-theories help us appreciate – is that the range of aspects which can have a therapeutic effect is both large and in many cases unintuitive. So for example: the mere number of pills (Blackwell et al. 1972; de Craen et al. 1999; Moerman 2000), the branding of pills (Branthwaite and Cooper 1981), whether one is given a pill or an injection (Amanzio et al. 2001), and the justified belief that one has undergone surgery (Cobb et al. 1959; Dimond et al. 1960), are just some examples of the different aspects of treatments that have been shown to have therapeutic consequences (See for more references: Koshi and Short 2007; Price et al. 2008; Stewart-Williams and Podd 2004). That is to say, more pills are better than fewer, branded pills are better than unbranded, injections work better than pills, and the justified belief that one has undergone surgery is itself sufficient for patients to improve. Other treatment aspects that have been shown to be therapeutically relevant include contextual factors such as verbal suggestions and the attitude, enthusiasm and behaviour of the healthcare team (Adler and Hammett 1973; Blasi et al. 2001; Kaptchuk 2002; Ong et al. 1995; Price et al. 2008) That is, all those verbal and non-verbal ways that patients and physicians interact to create a caring treatment context – a context which has demonstrable effects on patients' healing. Similarly the cognitive and emotional states of the patient are also aspects which affect patients' healing; for example Price et al (2008) and Stuart-Williams and Podd (2004) emphasise the role that a patient's expectations have been shown to play.

To perform a good placebo comparison we must ask questions about all the therapeutically relevant aspects of a treatment. As the experimental results above show, which aspects turn out to be relevant can be unintuitive. There is a

[5]Note aside that only some aspects of a treatment can be controlled for by randomisation (if that is, any can. See: Worrall 2007; 2010). In practice randomisation and the subsequent adjustment of baseline imbalances helps to minimise, for example, the influence of patients' differing expectations. However even in the ideal case randomisation would not help at all to solve the problem of treatment groups that were under the supervision of 'nasty doctor' on the one hand and 'nice doctor' on the other. The attitude and behaviour of the treating physician is something that we must seek to homogenise between groups through other means.

danger associated with calling certain objects or procedures ‘placebos’, in so far as this tempts us to forget to check they are genuinely ensuring the required identity between groups (See especially, in relation to pills and injections: Golomb 1995; Golomb et al. 2010).

If we want to know the efficacy of an extra 5mg of drug, on top of 10mg we can perform a placebo comparison, which compares two groups identical but for the fact that one receives 10mg of a drug and the other 15mg. When we follow the logic of placebo comparison, we see that it is a matter of how we set up certain features of the placebo group that matters, not what particular objects or procedures are employed. Sometimes placebo comparison may involve a placebo group which receives a pill containing 10mg of a drug as a control, because we are interested in the efficacy of a marginal 5mg above this. At other times placebo comparison may involve a placebo group which receives a pill containing only sugar as a control, because we are interested in the efficacy of a drug above the efficacy of pill-receiving. Both warrant being called placebo comparisons. There is no distinction worth making between the two that would make one a placebo comparison, and the other not.

Conclusion

I claim that the logic of placebo comparison amounts to nothing more than the idea that efficacy is attributed to some aspect of a treatment on the basis of differential effects between two groups which are identical in all but that respect. The properties that some set of objects or procedures will need to possess to ensure that some comparison is a genuine placebo comparison will depend entirely on the details of the aspect of the treatment being investigated. For the reason that how we achieve identity in all but one respect between groups will obviously differ according to the nature of the treatment and the aspect of interest.

These points are almost too obvious to note: it is the logic of placebo comparison that dictates the nature of the controls to be used when we set out to measure efficacy. The implications of this are less readily acknowledged: there is no sense besides arbitrary stipulation in calling an object or procedure, which in certain circumstances can do some of that controlling work, a ‘placebo’. Nor is there anything to be gained from distinguishing comparisons on the basis of whether they include objects or procedures that generate therapeutic responses by virtue of the meaning attached to them; that distinction is no more helpful

when measuring efficacy (that is, following the logic of – what I call – placebo comparison) than a distinction between trials that do or do not include objects that generate therapeutic responses through the renin-angiotensin system.

At best, the most meaningful thing we might say, while still using the term ‘placebo’, is that one group is the placebo group, while the other is the treatment group. That doesn’t get us very far however. We know that the properties of the placebo group are likely to differ depending on the nature of the treatment and the aspect of that treatment being investigated. We ask questions about those particular aspects of a treatment individually, such as whether the treatment is branded in both groups, what colour the pills are in both groups, whether patient’s were given the same information, or if control-pills contain any other therapeutically relevant chemicals – but that is not to ask anything about ‘placebos’, that is just to ask specific questions about certain aspects of a treatment. Once we know it is a placebo comparison we don’t need to invoke the term ‘placebo’ any longer. Rather, the meaningful questions to ask involve being specific about the details of the controls. ‘Placebo comparison’ only indicates a comparison with a particular epistemic aim; it indicates what kinds of specific questions to ask about the controls used.

I argue for abandoning the terms ‘placebo’ and ‘placebo effects’ because they serve no analytical purpose. We have a better view of what is going on in a placebo comparison if our descriptions don’t use those terms. In spite of this a medical researcher may object that the terms are perfectly functional, even if they are problematic. The point is that unless the terms are leading to clinically meaningful mistakes being made, then the argument above, in an important sense, does not matter. In response I would claim that the terms may well introduce practical problems. Talk of ‘placebos’ can tempt us to neglect questions about the adequacy of the placebo group to ensure the required identity to the treatment group[6]. To give one example: the credibility of trial results are often diminished where blinding has been unsuccessful (Rabkin et al, 1986). If identical-looking pills given to both groups differ, say, in taste or side-effects, then we have reason to worry about the success of the trial remaining blind. I admit that the extent to which this is a clinically meaningful problem is an empirical question; never the less, being explicit about how the control group

[6] Use of the term ‘placebo’ could also be important in a different way, if that usage created therapeutically relevant expectations in a patient. For example in a clinical context, through being told one is receiving ‘a placebo’; or in a research context, through being enrolled in a trial and told that one may be randomised to a placebo group. See for example: Kaptchuk et al, (2010) and Enck et al (2011).

was set-up is a matter of rigour. Talk of ‘placebos’ obscures legitimate questions about the specific details of the control group.

As explained in the introduction, I think the argument above supports the position advocated by Robin Nunn, who has argued that we should abandon the concept of ‘placebos’ and ‘placebo effects’ altogether. Perhaps the support is not total however, since I am happy to use the term ‘placebo comparison’, whereas Nunn is not. I use the term to refer to a particular kind of comparison; one with a specific logic behind it, namely, of making a comparison that singles out only one aspect of a treatment, and controls the rest. Understood in this way there is no imperative to use the term placebo comparison rather than some other (say, efficacy testing, and the logic of efficacy testing). So the difference between myself and Nunn is not substantive. The key point is that – what I would like to call – placebo comparison involves no commitments to ‘placebos’ or ‘placebo effects’.

Acknowledgements

Thanks to those at the 2010 ‘Progress in Medicine’ conference in Bristol who offered comments on a very early version of this paper, and thanks to Robin Nunn for comments on a much more recent version. Thanks also to the anonymous reviewers for their the helpful comments.

References

Adler HM, Hammett V (1973) The doctor-patient relationship revisited: An analysis of the placebo effect. Annals of Internal Medicine 78 (4):595-598

Amanzio M, Pollo A, Maggi G, Benedetti F (2001) Response variability to analgesics: a role for non-specific activation of endogenous opioids. Pain 90 (3):205-215

Blackwell B, Bloomfield S, Buncher C (1972) Demonstration to medical students of placebo responses and non-drug factors. The Lancet 299 (7763):1279-1282

Blasi ZD, Harkness E, Ernst E, Georgiou A, Kleijnen J (2001) Influence of context effects on health outcomes: a systematic review. The Lancet 357 (9258):757-762

Branthwaite A, Cooper P (1981) Analgesic effects of branding in treatment of headaches. BMJ 282 (6276):1576-1578

Cartwright N (2007) Are RCTs the Gold Standard? BioSocieties 2 (1):11-20

Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA (1959) An evaluation of internal-mammary-artery ligation by a double-blind technic. The New England Journal of Medicine 260 (22):1115-1118

de Craen AJM, Moerman DE, Heisterkamp SH, Tytgat GNJ, Tijssen JGP, Kleijnen J (1999) Placebo effect in the treatment of duodenal ulcer. British journal of clinical pharmacology 48 (6):853-860

Dimond EG, Kittle CF, Crockett JE (1960) Comparison of internal mammary artery ligation and sham operation for angina pectoris. The American Journal of Cardiology 5:483-486

Ernst E, Resch K-L (1995) Concept of true and perceived placebo effects. BMJ 311:551-553

Enck P, Klosterhalfen S, Weimer K, Horing B, Zipfel S. (2011) The placebo response in clinical trials: more questions than answers. Philosophical Transaction of the Royal Society B 366 (1572):1889-1895.

Golomb B (1995) Paradox of placebo effect. Nature 375:530

Golomb B, Erickson L, Koperski S, Sack D, Enkin M, Howick J (2010) What's in placebos: who knows? Analysis of randomized, controlled trials. Annals of Internal Medicine 153 (8):532

Gøtzsche PC (1994) Is there logic in the placebo? Lancet 344 (8927):925-926

Gøtzsche PC (1995) Placebo effects. Concept of placebo should be discarded. BMJ 311 (7020):1640

Grünbaum A (1981) The placebo concept. Behaviour research and therapy 19 (2):157

Grünbaum A (1991) The Placebo Concept in Medicine and Psychiatry. In: Cicchetti D, Grove WM (eds) Thinking Clearly about Psychology, Volume I: Matter of Public Interest. University of Minnesota Press, Oxford,

Hahn RA, Kleinman A (1983) Belief as pathogen, belief as medicine: "Voodoo death" and the" placebo phenomenon" in anthropological perspective. Medical Anthropology Quarterly 14 (4):3-19

Hill A (1951) The Clinical Trial. British Medical Bulletin 7 (4):278-282

Howick J (2009) Questioning the methodologic superiority of ‘placebo’over ‘active’controlled trials. The American Journal of Bioethics 9 (9):34-48

Hrobjartsson A, Gøtzsche PC (2001) Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. New England Journal of Medicine 344 (21):1594-1602

Hrobjartsson A, Gøtzsche PC (2004) Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. Journal of Internal Medicine 256 (2):91-100

Kaptchuk TJ (2002) The placebo effect in alternative medicine: can the performance of a healing ritual have clinical significance? Annals of Internal Medicine 136 (11):817-825

Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, et al. (2010) Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome. PLoS ONE 5(12): e15591

Kirmayer LJ (2004) The cultural diversity of healing: meaning, metaphor and mechanism. British Medical Bulletin 69 (1):33-48

Koshi E, Short C (2007) Placebo theory and its implications for research and clinical practice: a review of the recent literature. Pain Practice 7 (1):4-20

Miller FG, Kaptchuk TJ (2008) The power of context: reconceptualizing the placebo effect. Journal of the Royal Society of Medicine 101:222-225

Moerman D, Benoist J, Brody E, Giovannini M, Gracia M, Hall E, Heggenhougen H, Jonas D, Kearney M, Kedenburg D (1979) Anthropology of Symbolic Healing [and Comments and Reply]. Current Anthropology 20 (1):59-80

Moerman DE (2000) Cultural variations in the placebo effect: ulcers, anxiety, and blood pressure. Medical Anthropology Quarterly (New Series) 14 (1):51-72

Moerman DE (2002) Meaning, Medicine, and the 'Placebo Effect'. Cambridge University Press, Cambridge

Moerman DE, Jonas WB (2002) Deconstructing the placebo effect and finding the meaning response. Annals of Internal Medicine 136 (6):471-476

Nunn R (2009a) It's time to put the placebo out of our misery. BMJ 338 (apr20 2):b1568

Nunn R (2009b) Preparing for a Post-Placebo Paradigm: Ethics and Choice of Control in Clinical Trials. The American Journal of Bioethics 9 (9):51-52

Ong L, De Haes J, Hoos A, Lammes F (1995) Doctor-patient communication: a review of the literature. Social Science & Medicine 40 (7):903-918

Papakostas YG, Daras MD (2001) Placebos, placebo effect, and the response to the healing situation: The evolution of a concept. Epilepsia 42 (12):1614-1625

Price DD, Finniss DG, Benedetti F (2008) A Comprehensive Review of the Placebo Effect: Recent Advances and Current Thought. Annual Review of Psychology 59:565-590

Price L (1984) Art, science, faith and medicine: the implications of the placebo effect. Sociology of health & illness 6 (1):61-73

Rabkin JG, Markowitz JS, Stewart J, McGrath P, Harrison W, Quitkin FM, Klein DF (1986) How blind is blind? Assessment of patient and doctor medication guesses in a placebo-controlled trial of imipramine and phenelzine. Psychiatry Research 19:75–86

Shapiro AK (1964) A historic and heuristic definition of the placebo. Psychiatry 27:52-58

Shapiro AK (1968) Semantics of the placebo. Psychiatric Quarterly 42 (4):653-695

Shapiro AK, Shapiro E (1997a) The Placebo: Is It Much Ado About Nothing? In: Harrington A (ed) The Placebo Effect: An Interdisciplinary Exploration. Harvard University Press, London,

Shapiro AK, Shapiro E (1997b) The Powerful Placebo. Johns Hopkins University Press, London

Stein HF (1983)... On Placebos. To Cure, to Control, to Please: Medicine after the Demise of" The Placebo". Medical Anthropology Quarterly:4-17

Stewart-Williams S, Podd J (2004) The placebo effect: dissolving the expectancy versus conditioning debate. Psychological Bulletin 130 (2):324-340

Temple R, Ellenberg S (2000) Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: ethical and scientific issues. Annals of Internal Medicine 133 (6):455-463

Thompson J, Ritenbaugh C, Nichter M (2009) Reconsidering the Placebo Response from a Broad Anthropological Perspective. Culture, Medicine and Psychiatry 33 (1):112-152

Worrall J (2007) Evidence in Medicine and Evidence-Based Medicine. Philosophy Compass 2 (6):981-1022

Worrall J (2010) Evidence: philosophy of science meets medicine. Journal of Evaluation in Clinical Practice 16 (2):356-362